PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 34, Fasc. 1 (2014), pp. 1–22

PERSISTENCE PROBABILITIES FOR A BRIDGE OF AN INTEGRATED SIMPLE RANDOM WALK

BY

FRANK AURZADA* (DARMSTADT), STEFFEN DEREICH (MÜNSTER),
AND MIKHAIL LIFSHITS** (ST. PETERSBURG AND LINKÖPING)

Abstract. We prove that an integrated simple random walk, where random walk and integrated random walk are conditioned to return to zero, has asymptotic probability $n^{-1/2}$ to stay positive. This question is motivated by random polymer models and proves a conjecture by Caravenna and Deuschel.

2000 AMS Mathematics Subject Classification: Primary: 60G50; Secondary: 60F99.

Key words and phrases: Integrated random walk, local limit theorem, persistence probability.

1. INTRODUCTION AND MAIN RESULT

1.1. Introduction. This paper considers the *persistence probability*, i.e. the evaluation of

$$\mathbb{P}(A_1 \geqslant 0, \ldots, A_n \geqslant 0) \approx n^{-\theta}, \quad n \to \infty,$$

where A is some stochastic process and the number θ is called a *persistence exponent*. Here and below $f(n) \approx g(n)$ means that $f(n)/g(n)$ is bounded away from zero and infinity for large n. The problem is also sometimes called a *one-sided exit problem* or *survival probability problem*. Problems of this type have experienced quite some recent attention, see, e.g., [18], [3], [12], [1], [2], [19], [5], [6], [13], and the recent survey paper [4].

These probabilities have a couple of applications to problems in theoretical physics as well as to other questions in probability. We refer to the aforementioned survey [4] and to a survey article on the related physics literature [16] and its updated version [8] for details.

---

* Supported by the DFG Emmy Noether Programme.

** Supported by grants RFBR 13-01-00172, NSh-4472.2010.1, SPbSU 6.38.672.2013.

The particular problem that we treat here is motivated by a connection to random polymer models, see Section 1.5 in [10], also see [11]. Here, A will be an integrated simple random walk, where the pair of random walk and integrated random walk is conditioned to return to the origin. This is supposed to model a polymer chain with Laplace interaction and zero boundary conditions.

Let us be more precise and introduce the relevant notation. Let $(X_i)_{i\in\mathbb{N}}$ be a sequence of independent symmetric Bernoulli random variables ($\mathbb{P}(X_i = \pm 1) = 1/2$). We consider the simple random walk $S_n := \sum_{i=1}^n X_i$ and the respective integrated random walk $A_n := \sum_{i=1}^n S_i$ for $n \in \mathbb{N}_0 := \{0,1,\dots\}$. Note that the paired process $(S_n, A_n)_{n\in\mathbb{N}_0}$ is Markovian and one can easily check that it can return to $(0,0)$ only at the times $4n$, $n \geqslant 1$. Our main theorem is as follows.

THEOREM 1.1. *When $n \to \infty$ we have*

$$\mathbb{P}(A_1 \geqslant 0, \dots, A_{4n} \geqslant 0 \,|\, A_{4n} = S_{4n} = 0) \approx n^{-1/2}. \tag{1.1}$$

This proves the conjecture by Caravenna and Deuschel (see [10], (1.22) on p. 2396) for the case of the simple random walk.

Let us give a couple of remarks. The unconditioned probability has also been subject to a number of studies: The first is due to Sinaĭ [17] who showed that, with the notation above,

$$\mathbb{P}(A_1 \geqslant 0, \dots, A_n \geqslant 0) \approx n^{-1/4}. \tag{1.2}$$

This result was subsequently refined by [18], [3], [12], [19] to the end that

$$\mathbb{P}(A_1 \geqslant 0, \dots, A_{4n} \geqslant 0) \sim c n^{-1/4};$$

see [19], Theorem 1, which extends to other types of random walks. Here and below $f(n) \sim g(n)$ means that $f(n)/g(n) \to 1$ as $n \to \infty$.

We remark that the result in Theorem 1.1 is in contrast to conditioning only on $S_{4n} = 0$, where the rate is again

$$\mathbb{P}(A_1 \geqslant 0, \dots, A_{4n} \geqslant 0 \mid S_{4n} = 0) \approx n^{-1/4},$$

see [19], Proposition 1, where our problem is also mentioned. That is, conditioning on $S_{4n} = 0$ does not change the rate from the unconditioned situation.

The remainder of this paper is structured as follows: The next subsection gives an overview of the method of proof. In Section 2 we derive a local limit theorem for the (unconditioned) process (S_n, A_n), which we could not locate in the literature. A similar local limit theorem was proved in [10] (see Proposition 2.3 there) for random walks with X_1 having a continuous distribution and under an appropriate integrability assumption. In Section 3, we show some scaling properties of the process (S_n, A_n) under the condition $A_1 \geqslant 0, \dots, A_n \geqslant 0$. Further, Section 4 contains a CLT for the process (S_n, A_n) "pinned" at some final value: $(S_n, A_n) = (p_n, q_n)$

with $(n^{-1/2}p_n, n^{-3/2}q_n) \to (p,q)$; we show that the suitably scaled law of this process converges to the law of a pair of Brownian motion and integrated Brownian motion, conditioned to end at (p,q). This result may be of independent interest. Finally, Section 5 contains the proof of the main result.

1.2. Overview of the proof and notation. Throughout we use the following notation:

$$\Omega_n^+ := \{A_j \geqslant 0, 1 \leqslant j \leqslant n\}.$$

We let

$$D_n := \left\{ \ell = (\ell_1, \ell_2) \in \mathbb{Z}^2 : \ell_1 = n \ (\mathrm{mod}\ 2),\ \ell_2 = \frac{n(n+1)}{2} \ (\mathrm{mod}\ 2) \right\} \tag{1.3}$$

denote the set of all possible values of (S_n, A_n).

We will use the notation of an adjoint process which is gained via time reversion. Depending on parameters $N \in \mathbb{N}$ and $(s_N, a_N) \in \mathbb{Z}^2$ we define the adjoint process $(\bar{S}_n^{(N)}, \bar{A}_n^{(N)})_{n=0,\ldots,N}$ via

$$\bar{S}_0^{(N)} = s_N \quad \text{and} \quad \bar{A}_0^{(N)} = a_N$$

and the equations

$$\bar{S}_{n+1}^{(N)} = \bar{S}_n^{(N)} - X_{N-n} \quad \text{and} \quad \bar{A}_{n+1}^{(N)} = \bar{A}_n^{(N)} - \bar{S}_n^{(N)}.$$

By construction, one has for $n, m \in \{0, \ldots, N\}$

$$\{(\bar{S}_{N-n}^{(N)}, \bar{A}_{N-n}^{(N)}) = (S_n, A_n)\} = \{(\bar{S}_{N-m}^{(N)}, \bar{A}_{N-m}^{(N)}) = (S_m, A_m)\}, \tag{1.4}$$

meaning that the time reversed adjoint process and the original process either agree or disagree for all times $n = 0, \ldots, N$. In particular, the event of accordance is equal to $\{(\bar{S}_N^{(N)}, \bar{A}_N^{(N)}) = (0,0)\}$ or to $\{(S_N, A_N) = (s_N, a_N)\}$. The process $(\bar{S}_n^{(N)}, \bar{A}_n^{(N)})$ is Markovian and the definition can be extended in a canonical way to the time index $\mathbb{N}_0$.

The strategy of the proof is as follows: We partition the time frame $4n$ into three periods: in the first n steps, the process is observed under the conditioning. The same is done in the last n steps. Then in the middle $2n$ time steps one has to use (upper bound) or ensure (lower bound) that the two ends meet.

For the upper bound, we observe that, on Ω_{4n}^+ and with $A_{4n} = S_{4n} = 0$, one has Ω_n^+ and the same condition for the adjoint process on the final n time steps. Conditioning on the first and last n time steps, in the middle piece – consisting of $2n$ time steps – one again observes a pair of random walk and integrated random walk, however, starting and terminating at certain values. The probability of this pair starting and ending at certain values is governed by a local limit theorem.

The lower bound is more complex. Here one has to ensure that during the first (and last, respectively) n steps one ends with the pair (S_n, A_n) in a target zone $[an^{1/2}, bn^{1/2}] \times [an^{3/2}, bn^{3/2}]$, where a, b are appropriate positive constants. This can be shown by analyzing the scaling of S and A under the conditioning. To ensure that both ends meet, we prove a CLT for the process (S_n, A_n) which is "pinned" at the beginning and at the end (by values that are in the above target zone). This CLT helps us to transfer the question of positivity of the second component to the same question for the limiting process (that is, Brownian motion and its integrated counterpart), where the question of positivity is easily solved.

The local limit theorem just mentioned is stated and proved in Section 2. Then, in Section 3, we prove the results concerning the scaling of A and S under the conditioning needed in the proof of the lower bound. Further, the CLT for pinned process is formulated and proved in Section 4. Finally, in Section 5 we give the proof of our main theorem.

2. LOCAL LIMIT THEOREM FOR (S_n, A_n)

In order to state our limit theorem, we will need the notion of Brownian motion $B = (B_t)_{t\geqslant 0}$ and integrated Brownian motion $I = (I_t)_{t\geqslant 0}$ defined as

$$I_t := \int_0^t B_s\, ds.$$

Note that the paired process $\Gamma = (\Gamma_t)_{t\geqslant 0} = (B_t, I_t)_{t\geqslant 0}$ is a centered Gaussian Markov process and Γ_1 has the two-dimensional Lebesgue density

$$g(x,y) = \frac{\sqrt{3}}{\pi} \exp\{-2x^2 + 6xy - 6y^2\}, \tag{2.1}$$

which follows directly by calculating the covariance matrix of Γ.

A main ingredient of our proofs will be the following local limit theorem for the simple random walk and its integrated version, which we could not locate in the literature.

PROPOSITION 2.1. *We have*

$$\lim_{n\to\infty} \sup_{\ell \in D_n} \left| \frac{n^2}{4} \mathbb{P}\big((S_n, A_n) = \ell\big) - g\big(\ell_1/\sqrt{n}, \ell_2/n^{3/2}\big) \right| = 0,$$

where g is defined in (2.1).

2.1. Proof of Proposition 2.1.

Proof. We start with a general representation of probabilities through the characteristic function. Let $Y \in \mathbb{Z}^2$ be an integer random vector and

$$f(t) = \mathbb{E}e^{i(t,Y)} = \sum_{k\in\mathbb{Z}^2} \exp\{i(t_1k_1 + t_2k_2)\}\mathbb{P}(Y = k), \qquad t = (t_1, t_2) \in \mathbb{R}^2,$$

its characteristic function. Then for any $\ell \in \mathbb{Z}^2$

$$\begin{aligned}
&\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(t)\exp\{-i(t_1\ell_1+t_2\ell_2)\}dt_1dt_2 \\
&\quad = \sum_{k\in\mathbb{Z}^2}\mathbb{P}(Y=k)\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\exp\big\{i\big(t_1(k_1-\ell_1)+t_2(k_2-\ell_2)\big)\big\}dt_1dt_2 \\
&\quad = \sum_{k\in\mathbb{Z}^2}\mathbb{P}(Y=k)\int_{-\pi}^{\pi}\exp\{it_1(k_1-\ell_1)\}dt_1\int_{-\pi}^{\pi}\exp\{it_2(k_2-\ell_2)\}dt_2 \\
&\quad = \sum_{k\in\mathbb{Z}^2}\mathbb{P}(Y=k)(2\pi)^2\mathbb{1}_{\{k_1=\ell_1\}}\mathbb{1}_{\{k_2=\ell_2\}} = (2\pi)^2\mathbb{P}(Y=\ell).
\end{aligned}$$

We conclude that

$$\mathbb{P}(Y=\ell) = (2\pi)^{-2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(t)\exp\{-i(t_1\ell_1+t_2\ell_2)\}dt_1dt_2. \tag{2.2}$$

When Y is symmetric, then one has

$$\mathbb{P}(Y=\ell) = (2\pi)^{-2}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(t)\cos(t_1\ell_1+t_2\ell_2)dt_1dt_2. \tag{2.3}$$

In our case $Y_n = (S_n, A_n) \stackrel{d}{=} \big(\sum_1^n X_j, \sum_1^n jX_j\big)$, where (X_j) are independent Bernoulli variables. Hence, the characteristic function of Y_n is

$$f_{Y_n}(t) = \prod_{j=1}^{n}\cos(t_1 + jt_2).$$

Let us discuss a periodicity property of the integrand in (2.3). For the characteristic function we have

$$\begin{aligned}
f_{Y_n}(t_1+\pi, t_2) &= (-1)^n f_{Y_n}(t_1,t_2), \\
f_{Y_n}(t_1, t_2+\pi) &= (-1)^{[n(n+1)]/2} f_{Y_n}(t_1,t_2),
\end{aligned}$$

while for the cosine part we have

$$\begin{aligned}
\cos\big((t_1+\pi)\ell_1+t_2\ell_2\big) &= (-1)^{\ell_1}\cos(t_1\ell_1+t_2\ell_2), \\
\cos\big(t_1\ell_1+(t_2+\pi)\ell_2\big) &= (-1)^{\ell_2}\cos(t_1\ell_1+t_2\ell_2).
\end{aligned}$$

Therefore, if $\ell \in D_n$, then the integrand in (2.3) is π-periodical with respect to both coordinates. In particular, it is equal to one at any point of the set

$$Q := \{(0,0), (\pi,0), (0,\pi), (\pi,\pi)\}.$$

On the other hand, it is trivial to see that $|f_{Y_n}(t)| < 1$ for any $t \notin Q$, $n \geqslant 2$.

What follows next is a “three-domain approach” of proving a CLT through characteristic functions: we divide the area of integration into three pieces: on the first piece the integrand is exponentially small, the second piece is not large enough to give any contribution, and on the third piece, the integrand can be approximated by the corresponding normal characteristic function, and thus gives the main contribution.

Let us start with the first piece. If we look at the integral in (2.3) and consider the integration domain as a torus, then for large n the integral is essentially accumulated in the small vicinity of the set Q.

To be more precise, let $0 < \varepsilon < 1$, define the distance $d(t,s) := |t_1 - s_1| + n|t_2 - s_2|$, and let $d(t,Q) := \inf_{s\in Q} d(t,s)$. One can show first that the integrand on the domain $T_1 := \{t :\ d(t,Q) \geqslant \varepsilon\}$ is uniformly (in t and ℓ) exponentially small, that is

$$\sup_{t:\, d(t,Q)\geqslant\varepsilon} |f_{Y_n}(t)| \leqslant \big(1 - h(\varepsilon)\big)^n$$

for some h depending on ε, see Lemma 2.2 below. Therefore, everything reduces to the domain $\overline{T_1} := \{t : d(t,Q) < \varepsilon\}$. By periodicity, we can only consider

$$\{t :\ d(t,0) \leqslant \varepsilon\} = \{t :\ |t_1| + n|t_2| \leqslant \varepsilon\}$$

and then multiply the result by $|Q| = 4$.

Next, the zone $\{t :\ d(t,0) < \varepsilon\}$ is further split into two zones, $T_{2,M} := \{t : M/\sqrt{n} \leqslant d(t,0) < \varepsilon\}$ and $T_{3,M} := \{t :\ d(t,0) < M/\sqrt{n}\}$. On $T_{2,M}$ we use, for ε small enough, the bound

$$\begin{aligned}|f_{Y_n}(t)| &= \prod_{j=1}^{n} |\cos(t_1 + jt_2)| \leqslant \prod_{j=1}^{n} \exp\{-(t_1 + jt_2)^2/2\} \\ &= \exp\Big\{ - \sum_{j=1}^{n} (t_1 + jt_2)^2/2\Big\} \leqslant \exp\{-c(nt_1^2 + n^3 t_2^2)\},\end{aligned}$$

where we used Lemma 2.1 (see below) in the last step. Therefore,

$$\begin{aligned}&\lim_{M\to\infty} n^2 \int\limits_{T_{2,M}} |f_{Y_n}(t)|dt \\ &\qquad \leqslant \lim_{M\to\infty} n^2 \int\limits_{\{|t_1|+n|t_2|\geqslant M/\sqrt{n}\}} \exp\{-c(nt_1^2 + n^3 t_2^2)\}dt \\ &\qquad = \lim_{M\to\infty} \int\limits_{\{|s_1|+|s_2|\geqslant M\}} \exp\{-c(s_1^2 + s_2^2)\}ds = 0\end{aligned}$$

for any $n \in \mathbb{N}$, having used the change of variables $s_1 = t_1\sqrt{n}$, $s_2 = t_2 n^{3/2}$.

On $T_{3,M}$ by Taylor expansion we can compare f_{Y_n} with the corresponding normal characteristic function. Namely,

$$\begin{aligned}
f_{Y_n}(t) &= \prod_{j=1}^{n} \cos(t_1 + jt_2) = \exp\big\{ \sum_{j=1}^{n} \ln\cos(t_1 + jt_2)\big\} \\
&= \exp\big\{ \sum_{j=1}^{n} \ln\big(1 - (t_1 + jt_2)^2/2 + O(n^{-2})\big)\big\} \\
&= \exp\Big\{ -\frac{1}{2}\big(\sum_{j=1}^{n} (t_1 + jt_2)^2 + O(n^{-1})\big)\Big\} \\
&= \exp\Big\{ -\frac{1}{2}\big(nt_1^2 + 2\sum_{j=1}^{n} jt_1t_2 + \sum_{j=1}^{n} j^2t_2^2 + O(n^{-1})\big)\Big\} \\
&= \exp\bigg\{ -\frac{1}{2}\bigg(nt_1^2 + n^2t_1t_2 + \frac{n^3}{3}t_2^2 + O(n^{-1})\bigg)\bigg\}.
\end{aligned}$$

Therefore, using the change of variables $s_1 = t_1\sqrt{n}, s_2 = t_2 n^{3/2}$ in the second step, we get

$$\begin{aligned}
&\frac{n^2}{(2\pi)^2} \int_{T_{3,M}} f_{Y_n}(t) \exp\{-i(t_1\ell_1 + t_2\ell_2)\}dt_1dt_2 \\
&= \frac{n^2}{(2\pi)^2} \int_{\{|t_1|+n|t_2|\leqslant M/\sqrt{n}\}} f_{Y_n}(t) \exp\{-i(t_1\ell_1 + t_2\ell_2)\}dt_1dt_2 \\
&\to \frac{1}{(2\pi)^2}\times \\
&\quad\times \int_{\{|s_1|+|s_2|\leqslant M\}} \exp\left\{-\frac{1}{2}\left(s_1^2 + s_1s_2 + \tfrac{1}{3}s_2^2\right)\right\} \exp\{-i(s_1L_1 + s_2L_2)\}ds_1ds_2
\end{aligned}$$

as $n \to \infty$, provided that $\ell_1 = [L_1\sqrt{n}]$, $\ell_2 = \big[L_2\, n^{3/2}\big]$, and the convergence is uniform over $L_1, L_2 \in (0,\infty)$. For large M the latter limit is close, uniformly over L_1, L_2, to

$$\begin{aligned}
&\frac{1}{(2\pi)^2} \iint \exp\left\{-\frac{1}{2}(s_1^2 + s_1s_2 + \tfrac{1}{3}s_2^2)\right\} \exp\{-i(s_1L_1 + s_2L_2)\}ds_1ds_2 \\
&= \frac{\sqrt{\det R}}{2\pi} \frac{1}{2\pi\sqrt{\det R}} \iint \exp\left\{-\frac{1}{2}(R^{-1}s, s)\right\} \exp\{-i(s_1L_1 + s_2L_2)\}ds_1ds_2 \\
&= \frac{\sqrt{\det R}}{2\pi} \exp\left\{-\frac{1}{2}(RL, L)\right\},
\end{aligned}$$

where

$$R^{-1} = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} \end{pmatrix}, \quad R = \begin{pmatrix} 4 & -6 \\ -6 & 12 \end{pmatrix}, \quad \det R = 12.$$

By recalling the factor $|Q| = 4$, we arrive at

$$\frac{4\sqrt{12}}{2\pi}\exp\left\{-\frac{1}{2}(RL, L)\right\} = \frac{4\sqrt{3}}{\pi}\exp\{-2L_1^2 + 6L_1L_2 - 6L_2^2\},$$

as required by the proposition. ■

2.2. Some auxiliary lemmas.

LEMMA 2.1. *There exists $c > 0$ such that for all n large enough and any $t_1, t_2 \in \mathbb{R}$*

$$\sum_{j=1}^{n}(t_1 + jt_2)^2 \geqslant c(nt_1^2 + n^3t_2^2).$$

Proof. There is no loss of generality to assume $t_1 = -1$ and $t := t_2 \geqslant 0$. Then we have to evaluate the function

$$\sum_{j=1}^{n}(jt-1)^2 = S_2t^2 - 2S_1t + n,$$

where $S_2 = S_2(n) = \sum_{j=1}^{n} j^2 \sim n^3/3$ and $S_1 = S_1(n) := \sum_{j=1}^{n} j \sim n^2/2$.

Consider the function

$$G(t) := \sum_{j=1}^{n}(jt-1)^2 = S_2t^2 - 2S_1t + n - c_1(n + n^3t^2).$$

This is a quadratic function with discriminant

$$4S_1^2 - 4(n - c_1n)(S_2 - c_1n^3) \sim 4n^4\Big(\tfrac{1}{4} - (1-c_1)\big(\tfrac{1}{3} - c_1\big)\Big) < 0$$

for c_1 chosen sufficiently small and n sufficiently large. Thus G is positive everywhere. ■

LEMMA 2.2. *For any $0 < \varepsilon < \pi/2$ there exists $h = h(\varepsilon) \in (0, 1)$ such that for any integer $n \geqslant 4$ we have*

$$\sup_{t:\, d(t,Q)\geqslant\varepsilon} |f_{Y_n}(t)| \leqslant \big(1 - h(\varepsilon)\big)^n.$$

Proof. Let $M(x) := \min_{k\in\mathbb{Z}} |x - k\pi|$. Take any $t = (t_1, t_2) \in [0, \pi]^2$ such that $d(t, Q) \geqslant \varepsilon$. The latter means that

$$M(t_1) + nM(t_2) \geqslant \varepsilon.$$

Then two cases are possible:

(1) $nM(t_2) \leqslant \varepsilon/2$.

Then we have $M(t_1) \geqslant \varepsilon/2$. Hence, for any $1 \leqslant j \leqslant n/2$ it is true that

$$\begin{aligned} M(t_1 + jt_2) &\geqslant M(t_1) - M(jt_2) \geqslant M(t_1) - jM(t_2) \\ &\geqslant M(t_1) - nM(t_2)/2 \geqslant \varepsilon/2 - \varepsilon/4 = \varepsilon/4, \end{aligned}$$

and we obtain the required estimate

$$|f_{Y_n}(t)| \leqslant \prod_{j=1}^{n/2} |\cos(t_1 + jt_2)| \leqslant [\cos(\varepsilon/4)]^{n/2}.$$

(2) $nM(t_2) \geqslant \varepsilon/2$.

By symmetry reasons, there is no loss of generality in assuming that $0 < t_2 \leqslant \pi/2$.

Let $\delta := \varepsilon/10$. Then

$$\frac{nt_2}{\delta} = \frac{nM(t_2)}{\delta} \geqslant \frac{\varepsilon/2}{\varepsilon/10} = 5.$$

Let us choose an integer $m \geqslant 0$ such that $mt_2 \leqslant \delta \leqslant (m+1)t_2$. Since $mt_2 \leqslant \delta \leqslant nt_2/5$, we have $m \leqslant n/5$.

Assume for a while that $t_2 \leqslant \pi/3$. We show now that for any $1 \leqslant j \leqslant n - 2(m+1)$ the inequalities

$$M(t_1 + jt_2) < \delta \quad \text{and} \quad M\big(t_1 + jt_2 + 2(m+1)t_2\big) < \delta$$

are incompatible. Indeed, let the first one be satisfied. Then for some $k \in \mathbb{Z}$ we have

$$\pi k - \delta < t_1 + jt_2 < \pi k + \delta.$$

It follows that

$$t_1 + jt_2 + 2(m+1)t_2 > (\pi k - \delta) + 2\delta = \pi k + \delta \tag{2.4}$$

but (using twice that $\varepsilon < 1 < \pi/2$)

$$\begin{aligned} t_1 + jt_2 + 2(m+1)t_2 &< (\pi k + \delta) + 2(\delta + t_2) \\ &= \pi k + 3\delta + 2t_2 \leqslant \pi k + 3\pi/20 + 2\pi/3 \\ &< \pi(k+1) - \pi/20 \leqslant \pi(k+1) - \delta. \end{aligned} \tag{2.5}$$

From (2.4) and the last line of (2.5) it follows that $M\big(t_1 + jt_2 + 2(m+1)t_2\big) > \delta$, and incompatibility is proved. This fact yields

$$\#\{j \leqslant n : M(t_1 + jt_2) \geqslant \delta\} \geqslant \frac{n - 2(m+1)}{2} \geqslant \frac{3n}{10} - 1,$$

and

$$|f_{Y_n}(t)| = \prod_{j=1}^{n} |\cos(t_1 + jt_2)| \leqslant [\cos(\delta)]^{3n/10-1},$$

which settles the assertion of the lemma.

For the remaining case $\pi/3 \leqslant t_2 \leqslant \pi/2$ it is immediate to see that the inequalities

$$M(t_1 + jt_2) < \delta \quad \text{and} \quad M\big(t_1 + (j+1)t_2\big) < \delta$$

are incompatible; we obtain

$$\#\{j \leqslant n : M(t_1 + jt_2) \geqslant \delta\} \geqslant \frac{n-1}{2},$$

and

$$|f_{Y_n}(t)| = \prod_{j=1}^{n} |\cos(t_1 + jt_2)| \leqslant [\cos(\delta)]^{(n-1)/2}. \quad \blacksquare$$

3. SCALING OF S AND A UNDER THE CONDITIONING

The purpose of this section is to show the following facts concerning the scaling of S and A under the condition of positivity of A.

PROPOSITION 3.1. *There is a constant $c > 0$ such that, for all $n \geqslant 1$,*

$$\mathbb{E}\left(|S_n| \;\middle|\; \Omega_n^+\right) \leqslant cn^{1/2}, \tag{3.1}$$

$$\mathbb{E}\left(|A_n| \;\middle|\; \Omega_n^+\right) \leqslant cn^{3/2}. \tag{3.2}$$

Proof. (1) We start with the proof of (3.1). First we show that it suffices to estimate $\mathbb{E}(S_n^+|\Omega_n^+)$, since

$$\mathbb{E}\left(|S_n|\middle|\Omega_n^+\right) = \mathbb{E}(S_n^+ + S_n^-|\Omega_n^+) \leqslant 2\mathbb{E}(S_n^+|\Omega_n^+). \tag{3.3}$$

Indeed, if $S_n < 0$ we let σ_0 denote the last visit to zero: $\sigma_0 := \max\{k \leqslant n : S_k = 0\}$. Consider the path transformation that is given by inverting the steps after σ_0. This transformation maps any path in Ω_n^+ with $S_n < 0$ to a path in Ω_n^+ with $S_n > 0$. It is a one-to-one transformation (however, not a bijection, since the image of a transformation of a path in Ω_n^+ with $S_n > 0$ does not have to be in Ω_n^+). Therefore,

$$\mathbb{P}(\Omega_n^+ \cap \{S_n = -k\}) \leqslant \mathbb{P}(\Omega_n^+ \cap \{S_n = k\}), \qquad k > 0,$$

and thus

$$\mathbb{E}(S_n^-|\Omega_n^+) \leqslant \mathbb{E}(S_n^+|\Omega_n^+).$$

It remains to estimate the latter expectation. For $0 \leqslant t \leqslant n$, let

$$\Omega_{n,t} := \{A_j \geqslant 0, 1 \leqslant j \leqslant t; S_t = 0; S_j > 0, t+1 \leqslant j \leqslant n\}. \tag{3.4}$$

Clearly,

$$\Omega_n^+ = (\Omega_n^+ \cap \{S_n < 0\}) \cup \bigcup_{t=0}^{n} \Omega_{n,t}.$$

Thus,

$$\begin{aligned}
\mathbb{E}(S_n^+ | \Omega_n^+) &= \frac{1}{\mathbb{P}(\Omega_n^+)} \, \mathbb{E} S_n^+ \sum_{t=0}^{n} \mathbb{1}_{\Omega_{n,t}} \\
&= \sum_{t=0}^{n} \frac{\mathbb{E} S_n^+ \mathbb{1}_{\Omega_{n,t}}}{\mathbb{P}(\Omega_{n,t})} \, \frac{\mathbb{P}(\Omega_{n,t})}{\mathbb{P}(\Omega_n^+)} \\
&\leqslant \max_{0 \leqslant t \leqslant n} \mathbb{E}(S_n^+ | \Omega_{n,t}) \sum_{t=0}^{n} \frac{\mathbb{P}(\Omega_{n,t})}{\mathbb{P}(\Omega_n^+)} \\
&\leqslant \max_{0 \leqslant t \leqslant n} \mathbb{E}(S_n^+ | \Omega_{n,t}).
\end{aligned} \tag{3.5}$$

By definition,

$$\mathbb{E}(S_n^+ | \Omega_{n,t}) = \sum_{k>0} \frac{\mathbb{P}(\Omega_{n,t} \cap \{S_n = k\}) k}{\mathbb{P}(\Omega_{n,t})}.$$

We can represent each of the events here as an intersection of two independent events, respectively:

$$\begin{aligned}
\Omega_{n,t} &= \Omega_t^+ \cap \{S_t = 0\} \cap \{S_j - S_t > 0, t+1 \leqslant j \leqslant n\}, \\
\Omega_{n,t} \cap \{S_n = k\} &= \Omega_t^+ \cap \{S_t = 0\} \cap \{S_j - S_t > 0, t+1 \leqslant j \leqslant n; S_n - S_t = k\}.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\mathbb{E}\left(S_n^+ | \Omega_{n,t}\right) \\
&= \sum_{k>0} \frac{\mathbb{P}(\Omega_t^+ \cap \{S_t = 0\}) \mathbb{P}(S_j - S_t > 0, t+1 \leqslant j \leqslant n; S_n - S_t = k) k}{\mathbb{P}(\Omega_t^+ \cap \{S_t = 0\}) \mathbb{P}(S_j - S_t > 0, t+1 \leqslant j \leqslant n)} \\
&= \sum_{k>0} \frac{\mathbb{P}\left(S_j - S_t > 0, t+1 \leqslant j \leqslant n; S_n - S_t = k\right) k}{\mathbb{P}\left(S_j - S_t > 0, t+1 \leqslant j \leqslant n\right)} \\
&= \sum_{k>0} \frac{\mathbb{P}\left(S_i > 0, 1 \leqslant i \leqslant n-t; S_{n-t} = k\right) k}{\mathbb{P}\left(S_i > 0, 1 \leqslant i \leqslant n-t\right)} \\
&= \mathbb{E}(S_{n-t}^+ | S_i > 0, 1 \leqslant i \leqslant n-t), \qquad 0 \leqslant t \leqslant n.
\end{aligned} \tag{3.6}$$

In order to evaluate the latter expectation we use a stopping time argument. Let $v := \inf\{k : S_k = -1\}$ and $v_n := \min(v, n)$. Then v_n is a bounded stopping time and we have

$$0 = \mathbb{E} S_{v_n} = \mathbb{E} S_n \mathbb{1}_{v>n} - \mathbb{P}(v \leqslant n).$$

Hence, $\mathbb{E} S_n \mathbb{1}_{v>n} = \mathbb{P}(v \leqslant n) \leqslant 1$.

On the other hand, we know (see [15], XII.8) that $\mathbb{P}(v > n) \approx n^{-1/2}$.

Finally, let us consider $S'_i := S_{i+1} - S_1$, $0 \leqslant i \leqslant n$, and let v' be the corresponding stopping time. Then

$$\begin{aligned}(3.7)\qquad \mathbb{E}(S^+_{n+1}|S_i > 0, 1 \leqslant i \leqslant n+1) &= 1 + \mathbb{E}(S'_n|S'_i \geqslant 0, 1 \leqslant i \leqslant n)\\ &= 1 + \mathbb{E}(S'_n|v' > n)\\ &= 1 + \frac{\mathbb{E}S'_n \mathbb{1}_{v'>n}}{\mathbb{P}(v' > n)}\\ &\leqslant 1 + \sqrt{n}/c \leqslant C'\sqrt{n}.\end{aligned}$$

Combining this with (3.5) and (3.6) gives

$$\mathbb{E}(S^+_n|\Omega^+_n) \leqslant c\sqrt{n}. \tag{3.8}$$

This and (3.3) show (3.1).

(2) We now prove (3.2). We start with some simple estimates:

$$|A_n| = \Big|\sum_{k\leqslant n} S_k\Big| \leqslant \sum_{k\leqslant n} |S_k| = \sum_{k\leqslant n} S^+_k + \sum_{k\leqslant n} S^-_k.$$

Moreover, on Ω^+_n we have

$$0 \leqslant A_n = \sum_{k\leqslant n} S_k = \sum_{k\leqslant n} S^+_k - \sum_{k\leqslant n} S^-_k.$$

Hence,

$$|A_n| \leqslant 2\sum_{k\leqslant n} S^+_k \leqslant 2n \max_{0\leqslant k\leqslant n} S_k.$$

It is now enough to prove that for any $R \in \mathbb{N}$ it is true that

$$\mathbb{P}(\{\max_{0\leqslant k\leqslant n} S_k \geqslant R\} \cap \Omega^+_n) \leqslant 2\mathbb{P}(\{S_n \geqslant R\} \cap \Omega^+_n), \tag{3.9}$$

because this leads to the desired

$$\mathbb{E}\big(|A_n|\big|\Omega^+_n\big) \leqslant 2n\mathbb{E}\big(\max_{0\leqslant k\leqslant n} S_k\big|\Omega^+_n\big) \leqslant 4n\mathbb{E}(S^+_n|\Omega^+_n) \leqslant 4n\cdot c\sqrt{n} = c'n^{3/2},$$

where we used (3.8) in the third step.

For proving (3.9) we will only use the monotonicity property of Ω^+_n: if x and y are two paths with $x \in \Omega^+_n$ and $y \geqslant x$ pointwise then $y \in \Omega^+_n$.

Fix $R \in \mathbb{N}$. For any $S \in \{\max_{0\leqslant k\leqslant n} S_k \geqslant R\} \cap \Omega^+_n$ the time

$$\sigma_R := \max\{k \leqslant n : S_k = R\} \leqslant n$$

is well defined. For any $S \in \{\max_{0\leqslant k\leqslant n} S_k \geqslant R\} \cap \Omega_n^+ \cap \{S_n < R\}$ define its transformation y by inverting steps starting from σ_R. This is a one-to-one transformation and we have the following properties of y: $S_k = y_k$ for $k \leqslant \sigma_R$ while $S_k < R < y_k$ for $\sigma_R < k \leqslant n$ (in particular, $y_n > R$). Hence $S \leqslant y$ pointwise. By monotonicity, $y \in \Omega_n^+$.

We infer that our transformation is a one-to-one embedding (as a side note: it is not a bijection, since the image of the transformation of a path in $\Omega_n^+ \cap \{S_n > R\}$ may be outside Ω_n^+):

$$\{\max_{0\leqslant k\leqslant n} S_k \geqslant R\} \cap \Omega_n^+ \cap \{S_n < R\} \to \Omega_n^+ \cap \{S_n > R\}.$$

Hence,

$$\mathbb{P}(\{\max_{0\leqslant k\leqslant n} S_k \geqslant R\} \cap \Omega_n^+ \cap \{S_n < R\}) \leqslant \mathbb{P}(\Omega_n^+ \cap \{S_n > R\}).$$

and (3.9) follows. ■

The following lemma is also concerned with the scaling of S and A. We show that the joint distribution of S_n and A_n (conditioned on Ω_n^+) is not concentrated near zero when $n \to \infty$.

LEMMA 3.1. *For $l, m, n \in \mathbb{N}$ with $l < n$ one has*

$$\mathbb{P}\big(S_n \geqslant m, A_n \geqslant (n-l)m \mid \Omega_n^+\big) \geqslant \mathbb{P}(S_l \geqslant 2m)\, \mathbb{P}(|S_{n-l}| \leqslant m).$$

In particular, for any constants $c_1, c_2 > 0$, there exists a strictly positive constant $\kappa = \kappa(c_1, c_2)$ such that for all sufficiently large n

$$\mathbb{P}(S_n \geqslant c_1 n^{1/2}, A_n \geqslant c_2 n^{3/2} \mid \Omega_n^+) \geqslant \kappa > 0.$$

P r o o f. First note that

$$\begin{aligned}(3.10)\quad & \mathbb{P}\big(S_n \geqslant m,\ A_n \geqslant (n-l)m \mid \Omega_n^+\big) \\ & \geqslant \mathbb{P}(S_l \geqslant 2m, S_i - S_l \geqslant -m \text{ for } i = l+1, \ldots, n \mid \Omega_n^+) \\ & \geqslant \frac{\mathbb{P}(\{S_l \geqslant 2m\} \cap \Omega_l^+ \cap \{S_i - S_l \geqslant -m \text{ for } i = l+1, \ldots, n\})}{\mathbb{P}(\Omega_n^+)},\end{aligned}$$

since for $k = l+1, \ldots, n$

$$\begin{aligned} A_k &= A_l + \sum_{i=l+1}^{k} (S_i - S_l) + (k-l) S_l \\ &\geqslant 0 + \sum_{i=l+1}^{k} (-m) + (k-l) S_l \\ &= (k-l)(S_l - m) \geqslant (k-l)m \geqslant 0. \end{aligned}$$

By independence, the last term in (3.10) equals

$$\begin{aligned}\frac{\mathbb{P}(\{S_l \geqslant 2m\} \cap \Omega_l^+)\,\mathbb{P}(\min_{i=1,\dots,n-l} S_i \geqslant -m)}{\mathbb{P}(\Omega_n^+)} \\ \geqslant \frac{\mathbb{P}(S_l \geqslant 2m)\,\mathbb{P}(\Omega_l^+)\,\mathbb{P}(\max_{i=1,\dots,n-l} S_i \leqslant m)}{\mathbb{P}(\Omega_n^+)} \\ \geqslant \mathbb{P}(S_l \geqslant 2m)\,\mathbb{P}(|S_{n-l}| \leqslant m),\end{aligned}$$

where, in the last step, we used the reflection principle as well as the fact that

$$\mathbb{P}(\{S_l \geqslant 2m\} \cap \Omega_l^+) \geqslant \mathbb{P}(S_l \geqslant 2m) \cdot \mathbb{P}(\Omega_l^+), \tag{3.11}$$

which means that the events are positively correlated: Recall that a family of random variables $(X_i)_{1 \leqslant i \leqslant l}$ is called *associated* if for any pair of bounded coordinatewise non-decreasing functions $f_1, f_2 : \mathbb{R}^l \to \mathbb{R}^1$ it is true that

$$\mathbb{E}\big(f_1(X_1, \dots, X_l) f_2(X_1, \dots, X_l)\big) \geqslant \mathbb{E} f_1(X_1, \dots, X_l) \mathbb{E} f_2(X_1, \dots, X_l).$$

See [9] for detailed account of the association property and its extensions. One only needs to know that any family of independent random variables is associated, cf. Theorem 1.8 in [9] due to [14]. Thus, (3.11) holds. ■

4. CLT FOR THE PINNED PROCESS

In this section, we prove a CLT for the pinned process $(S_n, A_n)_{n=1,\dots,N}$, when letting $N \in 2\mathbb{N}$ tend to infinity. By "pinning" we mean that the process is conditioned to arrive at a certain point, depending on N and scaling in N with the natural scaling of the process. We restrict attention to even numbers N for technical reasons, although the following theorem remains valid for general N.

We need some more notation. Recall that the set D_n was defined in (1.3). We describe the pinning via an $\mathbb{R}^2$-valued sequence $(\mathfrak{p}^{(N)})_{N \in 2\mathbb{N}}$ satisfying

$$(N^{1/2}\mathfrak{p}_1^{(N)}, N^{3/2}\mathfrak{p}_2^{(N)}) \in D_N \quad \text{and} \quad \lim_{N \to \infty} \mathfrak{p}^{(N)} = \mathfrak{p}$$

for a $\mathfrak{p} \in \mathbb{R}^2$. Further, as described in the introduction, we associate with the process $(S_n, A_n)_{n=0,\dots,N}$ the adjoint process $(\bar{S}_n^{(N)}, \bar{A}_n^{(N)})_{n=0,\dots,N}$ started at

$$\bar{S}_0^{(N)} = N^{1/2}\mathfrak{p}_1^{(N)} \quad \text{and} \quad \bar{A}_0^{(N)} = N^{3/2}\mathfrak{p}_2^{(N)}.$$

We will consider both original process $(S_n, A_n)_{n=0,\dots,N}$ and adjoint process $(\bar{S}_n^{(N)}, \bar{A}_n^{(N)})_{n=0,\dots,N}$ in their normalized versions: we set for $s \in \frac{1}{N}\mathbb{Z} \cap [0,1]$ and $N \in 2\mathbb{N}$

$$\Xi_s^{(N)} := (N^{-1/2} S_{Ns}, N^{-3/2} A_{Ns}) \quad \text{and} \quad \bar{\Xi}_s^{(N)} := (N^{-1/2} \bar{S}_{Ns}^{(N)}, N^{-3/2} \bar{A}_{Ns}^{(N)})$$

and apply a continuous piecewise linear interpolation between the breakpoints $s \in \frac{1}{N}\mathbb{Z} \cap [0,1]$ to obtain continuous processes $\Xi^{(N)} = (\Xi_s^{(N)})_{s\in[0,1]}$ and $\bar{\Xi}^{(N)} = (\bar{\Xi}_s^{(N)})_{s\in[0,1]}$.

Using this notation, the CLT reads as follows.

THEOREM 4.1. *One has*

$$\mathcal{L}(\Xi^{(N)} \,|\, \Xi_1^{(N)} = \mathfrak{p}^{(N)}) \Rightarrow \mathcal{L}(\Gamma \,|\, \Gamma_1 = \mathfrak{p}),$$

where $\Gamma = (\Gamma_t)_{t\in[0,1]} = (B_t, I_t)_{t\in[0,1]}$.

REMARK 4.1. *We remark that the theorem remains valid when choosing different starting points for the Markov process* $(S_n, A_n)_{n\in\mathbb{N}}$. *Suppose that it is started in* $(s^{(N)}, a^{(N)})$ *such that the limit*

$$\mathfrak{s} = \lim_{N\to\infty} (N^{-1/2} s^{(N)}, N^{-3/2} a^{(N)})$$

exists. Now assuming that the pinning is done on non-null events, *one gets*

$$\mathcal{L}^{(s^{(N)},a^{(N)})}(\Xi^{(N)} \,|\, \Xi_1^{(N)}) \Rightarrow \mathcal{L}^{\mathfrak{s}}(\Gamma \,|\, \Gamma_1 = \mathfrak{p}).$$

Here the right-hand side denotes the law of integrated Brownian motion started in $\mathfrak{s}$. *The statement is straightforwardly obtained by using the fact that*

$$(\tilde{S}_n, \tilde{A}_n) := (S_n + s, A_n + a + ns)_{n\in\mathbb{N}}$$

has under $\mathbb{P}^{(0,0)}$ *the same distribution as* (S_n, A_n) *under* $\mathbb{P}^{(s,a)}$, *and the analogous property for the process* Γ.

In order to prove Theorem 4.1, we show tightness and convergence of finite-dimensional distributions for the conditioned distributions.

Proof of Theorem 4.1. *Tightness*. We prove that the sequence of conditional distributions $\mathcal{L}(\Xi^{(N)} \,|\, \Xi_1^{(N)} = \mathfrak{p}^{(N)})$ on $C\big([0,1], \mathbb{R}^2\big)$ is tight.

Let $\varepsilon > 0$ and fix compact sets K_1, K_2 in $C\big([0, \frac{1}{2}], \mathbb{R}^2\big)$ with

$$\mathbb{P}(\Xi^{(N)} \in K_1) \geqslant 1 - \varepsilon \quad \text{and} \quad \mathbb{P}(\bar{\Xi}^{(N)} \in K_2) \geqslant 1 - \varepsilon$$

for all $N \in 2\mathbb{N}$, where the processes are to be considered on the time interval $\big[0, \frac{1}{2}\big]$. Such compact sets exist by Donsker's invariance principle (see, e.g., [7]). Now let $K \subset C[0,1]$ be the set of continuous functions $f : [0,1] \to \mathbb{R}^2$ with

$$\big(f(t) : t \in \big[0, \tfrac{1}{2}\big]\big) \in K_1 \quad \text{and} \quad \big(f(1-t) : t \in \big[0, \tfrac{1}{2}\big]\big) \in K_2.$$

It is obviously compact in $C([0,1], \mathbb{R}^2)$. By the definition of the adjoint process (see (1.4)), one has

$$\{\Xi_1^{(N)} = \mathfrak{p}^{(N)}\} = \{\Xi_{1/2}^{(N)} = \bar{\Xi}_{1/2}^{(N)}\} = \{\bar{\Xi}_1^{(N)} = 0\},$$

so that

$$\begin{aligned}&\mathbb{P}(\Xi^{(N)} \notin K \,|\, \Xi_1^{(N)} = \mathfrak{p}^{(N)})\\ &\qquad \leqslant \mathbb{P}(\Xi^{(N)} \notin K_1 \,|\, \Xi_{1/2}^{(N)} = \bar{\Xi}_{1/2}^{(N)}) + \mathbb{P}(\bar{\Xi}^{(N)} \notin K_2 \,|\, \Xi_{1/2}^{(N)} = \bar{\Xi}_{1/2}^{(N)}).\end{aligned}$$

To obtain an upper bound for the first term, we observe that $(\mathbb{1}\{\Xi^{(N)} \notin K_1\}, \Xi_{1/2}^{(N)})$ and $\bar{\Xi}_{1/2}^{(N)}$ are independent, which implies that

$$\mathbb{P}(\Xi^{(N)} \notin K_1, \Xi_{1/2}^{(N)} = \bar{\Xi}_{1/2}^{(N)}) = \sum_z \mathbb{P}(\Xi^{(N)} \notin K_1, \Xi_{1/2}^{(N)} = z)\, \mathbb{P}(\bar{\Xi}_{1/2}^{(N)} = z).$$

By the local central limit theorem, the weights $\mathbb{P}(\bar{\Xi}_{1/2}^{(N)} = z)$ are uniformly bounded by a constant multiple of N^{-2} so that

$$\mathbb{P}(\Xi^{(N)} \notin K_1, \Xi_{1/2}^{(N)} = \bar{\Xi}_{1/2}^{(N)}) \leqslant C_1 \frac{1}{N^2} \mathbb{P}(\Xi^{(N)} \notin K_1) \leqslant C_1 \varepsilon \frac{1}{N^2}$$

for a universal constant C_1. Analogously, one concludes that

$$\mathbb{P}(\bar{\Xi}^{(N)} \notin K_2, \bar{\Xi}_{1/2}^{(N)} = \Xi_{1/2}^{(N)}) \leqslant C_2 \frac{1}{N^2} \mathbb{P}(\bar{\Xi}^{(N)} \notin K_1) \leqslant C_2 \varepsilon \frac{1}{N^2}$$

for a universal constant C_2. Since by the local central limit theorem (Proposition 2.1 above) $\lim_{N\to\infty} N^2\, \mathbb{P}(\Xi_1^{(N)} = \mathfrak{p}^{(N)}) = C_3 > 0$, we conclude that

$$\limsup_{N\to\infty} \mathbb{P}(\Xi^{(N)} \notin K \,|\, \Xi_1^{(N)} = \mathfrak{p}^{(N)}) \leqslant \frac{C_1 + C_2}{C_3}\, \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, this proves tightness.

Convergence of finite-dimensional distributions. It remains to prove convergence of finite-dimensional marginals. For $t > 0$ we denote by $g_t : \mathbb{R}^2 \times \mathbb{R}^2 \to [0, \infty)$ the transition density of the Markov process (Γ_s) over an interval of length t. It is

$$g_t(u, v; x, y) = t^{-2}\, g\left(\frac{x - u}{\sqrt{t}}, \frac{y - v - tu}{t^{3/2}}\right).$$

Fix $m \in \mathbb{N}$, times $0 < t_1 < \ldots < t_m < 1$, and a continuous and bounded function $f : (\mathbb{R}^2)^m \to [0, \infty)$. We write $t_m^* = t_m^*(N) = \min(\mathbb{Z}/N) \cap [t_m, 1]$.

First we verify that, for arbitrary $\varepsilon > 0$, it follows that for sufficiently large $N \in 2\mathbb{N}$

$$\mathbb{P}(\Xi_1^{(N)} = \mathfrak{p}^{(N)} \,|\, \Xi_{t_m^*}^{(N)} = z) \geqslant \big(4 g_{1-t_m^*}(z, \mathfrak{p}) - \varepsilon\big) N^{-2} \tag{4.1}$$

for all z with $\mathbb{P}(\Xi_{t_m^*}^{(N)} = z) > 0$. The estimate is a consequence of the local limit theorem (Proposition 2.1): we let

$$\ell = \ell(N) = \big(\sqrt{N}\mathfrak{p}_1^{(N)}, N^{3/2}\mathfrak{p}_2^{(N)}\big)$$

along with $n = n(N) = (1 - t_m^*)N$ and $\zeta = \big(\sqrt{N} z_1, N^{3/2} z_2\big)$, where z is as before. For any $\varepsilon' > 0$, it follows that uniformly in the relevant z's we obtain for sufficiently large N

$$\begin{aligned}(4.2)\quad \mathbb{P}(\Xi_1^{(N)} = \mathfrak{p}^{(N)} \,|\, \Xi_{t_m^*}^{(N)} = z) &= \mathbb{P}^{\zeta}\big((S_n, A_n) = \ell\big)\\ &= \mathbb{P}\big((S_n, A_n) = (\ell_1 - \zeta_1, \ell_2 - \zeta_2 - n\zeta_1)\big)\\ &\geqslant 4n^{-2} g\Big(\frac{\ell_1 - \zeta_1}{\sqrt{n}}, \frac{\ell_2 - \zeta_2 - n\zeta_1}{n^{3/2}}\Big) - \varepsilon' n^{-2}\\ &= 4N^{-2} g_{1-t_m^*}(z_1, z_2; \mathfrak{p}_1^{(N)}, \mathfrak{p}_2^{(N)}) - \varepsilon' n^{-2}.\end{aligned}$$

Since $n(N)$ is of order N, we can choose for given $\varepsilon > 0$ a sufficiently small $\varepsilon' > 0$ such that (4.2) implies (4.1).

Hence, by the Markov property one has

$$\begin{aligned}&\mathbb{E}[f(\Xi_{t_1}^{(N)}, \dots, \Xi_{t_m}^{(N)})\, \mathbb{1}_{\{\Xi_1^{(N)} = \mathfrak{p}^{(N)}\}}]\\ &\qquad = \mathbb{E}\big[f(\Xi_{t_1}^{(N)}, \dots, \Xi_{t_m}^{(N)})\, \mathbb{E}[\mathbb{1}_{\{\Xi_1^{(N)} = \mathfrak{p}^{(N)}\}} | \mathcal{F}_{t_m^*}]\big]\\ &\qquad \geqslant 4N^{-2}\, \mathbb{E}[f(\Xi_{t_1}^{(N)}, \dots, \Xi_{t_m}^{(N)})\, g_{1-t_m^*}(\Xi_{t_m^*}^{(N)}, \mathfrak{p})] - \varepsilon C\, N^{-2},\end{aligned}$$

where $C = \|f\|_\infty$. Using the continuity of g together with the classical Donsker invariance principle [7], we arrive at

$$\begin{aligned}&\liminf_{n\to\infty} \tfrac{1}{4} N^2 \mathbb{E}[f(\Xi_{t_1}^{(N)}, \dots, \Xi_{t_l}^{(N)})\, \mathbb{1}_{\{\Xi_1^{(N)} = \mathfrak{p}^{(N)}\}}]\\ &\qquad\qquad \geqslant \mathbb{E}[f(\Gamma_{t_1}, \dots, \Gamma_{t_m})\, g_{1-t_m}(\Gamma_{t_m}, \mathfrak{p})] - \varepsilon C.\end{aligned}$$

Analogously, one proves the converse bound. Since $\varepsilon > 0$ is arbitrary, we get

$$\begin{aligned}\lim_{n\to\infty} \tfrac{1}{4} N^2\, \mathbb{E}[f(\Xi_{t_1}^{(N)}, \dots, \Xi_{t_m}^{(N)})\, \mathbb{1}_{\{\Xi_1^{(N)} = \mathfrak{p}^{(N)}\}}] &= \mathbb{E}[f(\Gamma_{t_1}, \dots, \Gamma_{t_m})\, g_{1-t_m}(\Gamma_{t_m}, \mathfrak{p})]\\ &= \int \dots \int f(z_1, \dots, z_l)\, g_{t_1}(0, z_1) \dots g_{1-t_m}(z_m, \mathfrak{p})\, dz_1 \dots dz_m.\end{aligned}$$

In particular,

$$\lim_{n\to\infty} \tfrac{1}{4} N^2\, \mathbb{E}[\mathbb{1}_{\{\Xi_1^{(N)} = \mathfrak{p}^{(N)}\}}] = g_1(0, \mathfrak{p}),$$

and putting the estimates together completes the proof. ∎

5. PROOF OF THE MAIN THEOREM

5.1. Proof of the upper bound. The purpose of this section is to prove the upper bound in (1.1). This will follow almost directly from a local limit theorem for (S_n, A_n) (Proposition 2.1).

Let us recall that

$$\Omega_n^+ = \{A_1 \geqslant 0, \dots, A_n \geqslant 0\} \in \sigma(X_1, \dots, X_n)$$

and let us define

$$\bar{\Omega}_n^+ = \{\bar{A}_1^{(4n)} \geqslant 0, \dots, \bar{A}_n^{(4n)} \geqslant 0\} \in \sigma(X_{3n+1}, \dots, X_{4n}),$$

where the adjoint process is started at $(0,0)$. Then due to the fact that $\bar{A}_k^{(4n)} = A_{4n-k}$ on $A_{4n} = S_{4n} = 0$ (see (1.4)) we have

$$\begin{aligned}\mathbb{P}\,(A_1 \geqslant 0, \dots, A_{4n} \geqslant 0, A_{4n} = S_{4n} = 0)&\\ &\leqslant \mathbb{P}(\Omega_n^+ \cap \bar{\Omega}_n^+ \cap \{A_{4n} = S_{4n} = 0\})\\ &= \mathbb{P}(A_{4n} = S_{4n} = 0 \,|\, \Omega_n^+ \cap \bar{\Omega}_n^+) \cdot \mathbb{P}(\Omega_n^+ \cap \bar{\Omega}_n^+).\end{aligned}$$

Clearly, Ω_n^+ and $\bar{\Omega}_n^+$ are independent. Further, the adjoint process started at $(0,0)$ has the same distribution as the original process, so that $\mathbb{P}(\Omega_n^+) = \mathbb{P}(\bar{\Omega}_n^+) \approx n^{-1/4}$, by Sinaĭ's result [17].

On the other hand, the event $\Omega_n^+ \cap \bar{\Omega}_n^+$ only concerns the random variables $X_1, \dots, X_n$ and $X_{3n+1}, \dots, X_{4n}$, so that

$$\begin{aligned}(5.1)\quad &\mathbb{P}(A_{4n} = S_{4n} = 0 \,|\, \Omega_n^+ \cap \bar{\Omega}_n^+)\\ &\leqslant \sup_{\substack{x_1,\dots,x_n,\\ x_{3n+1},\dots,x_{4n} \in \{-1,+1\}}} \mathbb{P}(A_{4n} = S_{4n} = 0 \,|X_1 = x_1, \dots, X_n = x_n,\\ &\qquad\qquad \text{and } X_{3n+1} = x_{3n+1}, \dots, X_{4n} = x_{4n}).\end{aligned}$$

Given the values for $X_1, \dots, X_n$ and $X_{3n+1}, \dots, X_{4n}$, the variables (S_i, A_i), $i \in \{n+1, \dots, 3n\}$, form another pair of simple random walk and its integrated counterpart, however, the pair is started at some different point: that is, for the vector (S_i, A_i), $i = n+1, \dots, 2n$, started at (k,l) has the distribution

$$\big(k + S_{i-n}, k(i-n-1) + l + A_{i-n}\big).$$

Thus, the quantity in (5.1) is bounded by

$$\sup_{a,s,a',s'} \mathbb{P}^{(s,a)}(A_{2n} = a', S_{2n} = s') = \sup_{a',s'} \mathbb{P}^{(0,0)}(A_{2n} = a', S_{2n} = s').$$

Now, the local limit theorem for (S_n, A_n) (Proposition 2.1) tells us that the latter probability is bounded by cn^{-2}.

Summing up and using once again the local limit theorem, we get

$$\begin{aligned}\mathbb{P}\,(A_1 \geqslant 0, \dots, A_{4n} \geqslant 0 \,|A_{4n} = S_{4n} = 0\,)&\\ &= \frac{\mathbb{P}\,(A_1 \geqslant 0, \dots, A_{4n} \geqslant 0, A_{4n} = S_{4n} = 0)}{\mathbb{P}\,(A_{4n} = S_{4n} = 0)}\\ &\leqslant \text{const} \cdot \frac{n^{-2} \cdot n^{-1/4} \cdot n^{-1/4}}{n^{-2}} \approx n^{-1/2},\end{aligned}$$

which completes the proof of the upper bound.

5.2. Proof of the lower bound. Here we prove the lower bound in (1.1).

First observe that Proposition 3.1, Lemma 3.1, and the Markov inequality show that there are constants $0 < a < b < \infty$ and $\kappa > 0$ such that, for all n,

$$\mathbb{P}(S_n \in [an^{1/2}, bn^{1/2}], A_n \in [an^{3/2}, bn^{3/2}] \mid \Omega_n^+) \geqslant \kappa > 0. \tag{5.2}$$

Using the Markov property of (S_n, A_n) we have

$$\begin{aligned}\mathbb{P}\big(\Omega_{4n}^+ \cap \{(S_{4n}, A_{4n}) = (0,0)\}\big) \\ &= \sum_{k,l}\sum_{k',l'} \mathbb{P}^{(0,0)}\big(\Omega_n^+ \cap \{(S_n, A_n) = (k,l)\}\big) \\ &\quad \times \mathbb{P}^{(k,l)}\big(\Omega_{2n}^+ \cap \{(S_{2n}, A_n) = (k',l')\}\big) \\ &\quad \times \mathbb{P}^{(k',l')}\big(\Omega_n^+ \cap \{(S_n, A_n) = (0,0)\}\big).\end{aligned} \tag{5.3}$$

We denote by $(\bar{S}_k^{(n)}, \bar{A}_k^{(n)})_{k=0,\dots,n}$ the adjoint process started in $(0,0)$ and consider

$$\bar{\Omega}_n^+ = \{\bar{A}_k^{(n)} \geqslant 0 \text{ for all } k = 0, \dots, n\}.$$

By (1.4), one has

$$\begin{aligned}\mathbb{P}^{(k',l')}\big(\Omega_n^+ \cap \{(S_n, A_n) = (0,0)\}\big) &= \mathbb{P}^{(k',l')}\big(\bar{\Omega}_n^+ \cap \{(\bar{S}_n^{(n)}, \bar{A}_n^{(n)}) = (k',l')\}\big) \\ &= \mathbb{P}^{(0,0)}\big(\Omega_n^+ \cap \{(S_n, A_{n-1}) = (-k',l')\}\big),\end{aligned}$$

where in the last step we used the fact that $\big((\bar{S}_m^{(n)}, \bar{A}_m^{(n)}) : m = 0, \dots, n\big)$ under $\mathbb{P}^{(k',l')}$ is distributed as $\big((-S_m, A_{m-1}) : m = 0, \dots, n\big)$ under $\mathbb{P}^{(0,0)}$ with target point (k',l'). Here we adopt the convention $A_{-1} = A_0 - S_0$. In order to compute a lower bound for (5.3), we can confine ourselves to summands where (k,l) and $(-k',l')$ are in $[an^{1/2}, bn^{1/2}] \times [an^{3/2}, bn^{3/2}]$, respectively, that is to

$$\mathcal{L}(n) = \{(k,l) \in D_n \mid (k,l) \in [an^{1/2}, bn^{1/2}] \times [an^{3/2}, bn^{3/2}]\}$$

and

$$\mathcal{R}(n) = \{(k',l') \in \tilde{D}_n \mid (-k',l') \in [an^{1/2}, bn^{1/2}] \times [an^{3/2}, bn^{3/2}]\},$$

respectively, where D_n is defined in (1.3) and

$$\tilde{D}_n := \left\{\ell = (\ell_1, \ell_2) \in \mathbb{Z}^2 : \ell_1 = n \,(\text{mod } 2),\ \ell_2 = \frac{(n-1)n}{2} \,(\text{mod } 2)\right\}.$$

We will prove below that

$$2\kappa' := \liminf_{n\to\infty} n^2 \inf_{\substack{(k,l)\in\mathcal{L}(n)\\(k',l')\in\mathcal{R}(n)}} \mathbb{P}^{(k,l)}\big(\Omega_{2n}^+ \cap \{(S_{2n}, A_{2n}) = (k',l')\}\big) > 0. \tag{5.4}$$

Then one estimates (5.3) by the product of $\kappa' n^{-2}$,

$$\sum_{(k,l)\in\mathcal{L}(n)} \mathbb{P}\big(\Omega_n^+ \cap \{(S_n, A_n) = (k,l)\}\big) \geqslant \mathbb{P}(\Omega_n^+)\kappa,$$

and

$$\sum_{(k',l')\in\mathcal{R}(n)} \mathbb{P}\big(\Omega_n^+ \cap \{(S_n, A_{n-1}) = (-k',l')\}\big) \geqslant \mathbb{P}(\Omega_n^+)\kappa,$$

where we used (5.2) in both cases. To see the second relation one needs the following additional argument in conjunction with (5.2): Let $0 < \varepsilon < b - a$. Then for all n large enough

$$\begin{aligned}&\mathbb{P}\big(S_n \in [an^{1/2}, bn^{1/2}], A_{n-1} \in [(a-\varepsilon)n^{3/2}, bn^{3/2}] \mid \Omega_n^+\big)\\ &\qquad \geqslant \mathbb{P}\big(S_n \in [an^{1/2}, bn^{1/2}], A_n \in [an^{3/2}, bn^{3/2}] \mid \Omega_n^+\big) \geqslant \kappa > 0;\end{aligned}$$

and we continue to work with $a - \varepsilon$ instead of a.

Thus, assuming (5.4), using the local limit theorem (Proposition 2.1) for the denominator, and Sinaĭ's result [17], we get

$$\begin{aligned}\mathbb{P}\big(\Omega_{4n}^+ \mid (S_{4n}, A_{4n}) = (0,0)\big) &= \frac{\mathbb{P}\big(\Omega_{4n}^+ \cap \{(S_{4n}, A_{4n}) = (0,0)\}\big)}{\mathbb{P}\big((S_{4n}, A_{4n}) = (0,0)\big)}\\ &\geqslant \frac{\kappa' n^{-2} \cdot \kappa c\, n^{-1/4} \cdot \kappa c\, n^{-1/4}}{c\,(4n)^{-2}} \approx n^{-1/2},\end{aligned}$$

which shows the assertion.

It remains to prove (5.4). We proceed with a proof by contradiction.

Assume that the $\liminf$ in (5.4) is zero. Then there exist an $\mathbb{N}$-valued sequence $(n_m)_{m\in\mathbb{N}}$ that tends to infinity and pairs $(k_m, l_m) \in \mathcal{L}(n_m)$ and $(k'_m, l'_m) \in \mathcal{R}(n_m)$ for $m \in \mathbb{N}$ such that

$$\lim_{m\to\infty} n_m^2\, \mathbb{P}^{(k_m,l_m)}\big(\Omega_{2n_m}^+ \cap \{(S_{2n_m}, A_{2n_m}) = (k'_m, l'_m)\}\big) = 0.$$

Without loss of generality we can assume that the limits

$$\begin{aligned}\mathfrak{s} &:= \lim_{m\to\infty} \big((2n_m)^{-1/2} k_m, (2n_m)^{-3/2} l_m\big) \in [a,b]^2,\\ \mathfrak{p} &:= \lim_{m\to\infty} \big((2n_m)^{-1/2} k'_m, (2n_m)^{-3/2} l'_m\big) \in [a,b]^2\end{aligned}$$

exist, since this is the case for at least one subsequence of (n_m). In order to apply the central limit theorem, we work with the continuous function

$$F : C[0,1]^2 \to [0,\infty),\ (z^1, z^2) \mapsto 1 \wedge \big(\inf_{t\in[0,1]} z_t^2\big)^+ \leqslant \mathbb{1}_{\{z_t^2 \geqslant 0, t\in[0,1]\}}.$$

Let
$$Z_t^m := (Z_t^{m,1}, Z_t^{m,2}) := \big((2n_m)^{-1/2}\, S_{2n_m t}, (2n_m)^{-3/2}\, A_{2n_m t}\big)$$
for $t \in \mathbb{N}_0/(2n_m)$ and extend the definition of Z^m between the points in $\mathbb{N}_0/(2n_m)$ linearly. Then

$$\begin{aligned}&\mathbb{P}^{(k_m,l_m)}\big(\Omega_{2n_m}^+ \cap \{(S_{2n_m}, A_{2n_m}) = (k'_m, l'_m)\}\big)\\ &\geqslant \mathbb{E}^{(k_m,l_m)}[F(Z) \mid (S_{2n_m}, A_{2n_m}) = (k'_m, l'_m)] \cdot \mathbb{P}\big((S_{2n_m}, A_{2n_m}) = (k'_m, l'_m)\big).\end{aligned}$$

Since $\mathbb{P}^{(k_m,l_m)}\big((S_{2n_m}, A_{2n_m}) = (k'_m, l'_m)\big) \geqslant c(2n_m)^{-2}$ for a positive constant c, a contradiction is achieved once we show that

$$\limsup_{m\to\infty} \mathbb{E}^{(k_m,l_m)}[F(Z) \mid (S_{2n_m}, A_{2n_m}) = (k'_m, l'_m)] > 0.$$

However, this follows directly from the local limit theorem (Proposition 2.1) and Remark 4.1. Indeed, the $\limsup$ is actually a limit and it is equal to

$$\mathbb{E}^{\mathfrak{s}}[F(\Gamma)|\Gamma_1 = \mathfrak{p}].$$

Furthermore, it is positive, since $\mathfrak{s}_2$ and $\mathfrak{p}_2$ are positive.

REMARK 5.1. After the first version of this paper had been made public in the preprint arXiv:1205.2895, Denisov and Wachtel announced in [13] an extension of our main theorem, cf. their Section 1.3. The formal proof is not given, however, their methods are – presumably – quite different from ours.

An anonymous referee pointed out that the strategy suggested in the present paper might also be valid for a much more general class of random walks. We leave the exploration of this valuable idea to subsequent publications.

Acknowledgments. We would like to thank the two anonymous referees for their careful reading and their suggestions.

REFERENCES

[1] F. Aurzada, *On the one-sided exit problem for fractional Brownian motion*, Electron. Comm. Probab. 16 (2011), pp. 392–404.

[2] F. Aurzada and C. Baumgarten, *Survival probabilities for weighted random walks*, ALEA Lat. Am. J. Probab. Math. Stat. 8 (2011), pp. 235–258.

[3] F. Aurzada and S. Dereich, *Universality of the asymptotics of the one-sided exit problem for integrated processes*, Ann. Inst. H. Poincaré Probab. Statist. 49 (2013), pp. 236–251.

[4] F. Aurzada and T. Simon, *Persistence probabilities & exponents*, preprint, arXiv: 1203.6554, 2012; Lévy Matters, Springer (to appear).

[5] C. Baumgarten, *Survival probabilities of some iterated processes*, preprint, arXiv: 1106.2999, 2011; Probab. Math. Statist. (to appear).

[6] C. Baumgarten, *Survival probabilities of autoregressive processes*, ESAIM Probab. Stat. 18 (2014), pp. 145–170.

[7] P. Billingsley, *Convergence of Probability Measures*, second edition, Wiley Ser. Probab. Stat., Wiley, New York 1999.
[8] A. J. Bray, S. N. Majumdar, and G. Schehr, *Persistence and first-passage properties in non-equilibrium systems*, Adv. in Phys. 62 (3) (2013), pp. 225–361.
[9] A. Bulinski and A. Shashkin, *Limit Theorems for Associated Random Fields and Related Systems*, Adv. Ser. Stat. Sci. Appl. Probab., Vol. 10, World Scientific, Hackensack, NJ, 2007.
[10] F. Caravenna and J.-D. Deuschel, *Pinning and wetting transition for* $(1{+}1)$*-dimensional fields with Laplacian interaction*, Ann. Probab. 36 (6) (2008), pp. 2388–2433.
[11] F. Caravenna and J.-D. Deuschel, *Scaling limits of* $(1+1)$*-dimensional pinning models with Laplacian interaction*, Ann. Probab. 37 (3) (2009), pp. 903–945.
[12] A. Dembo, J. Ding, and F. Gao, *Persistence of iterated partial sums*, Ann. Inst. H. Poincaré Probab. Statist. 49 (3) (2013), pp. 873–884.
[13] D. Denisov and V. Wachtel, *Exit times for integrated random walks*, preprint, arXiv:1207.2270, 2012; Ann. Inst. H. Poincaré Probab. Statist. (to appear).
[14] J. D. Esary, F. Proschan, and D. W. Walkup, *Association of random variables, with applications*, Ann. Math. Statist. 38 (5) (1967), pp. 1466–1474.
[15] W. Feller, *An Introduction to Probability Theory and Its Applications. Vol. II*, second edition, Wiley, New York 1971.
[16] S. N. Majumdar, *Persistence in nonequilibrium systems*, Current Sci. 77 (3) (1999), pp. 370–375.
[17] Ya. G. Sinaĭ, *Distribution of some functionals of the integral of a random walk*, Teoret. Mat. Fiz. 90 (3) (1992), pp. 323–353.
[18] V. Vysotsky, *On the probability that integrated random walks stay positive*, Stochastic Process. Appl. 120 (7) (2010), pp. 1178–1193.
[19] V. Vysotsky, *Positivity of integrated random walks*, Ann. Inst. H. Poincaré Probab. Statist. 50 (1) (2014), pp. 195–213.

Frank Aurzada
Technische Universität Darmstadt
FB Mathematik, AG Stochastik
Schlossgartenstr. 7
64289 Darmstadt, Germany
E-mail: f.aurzada@tu-braunschweig.de

Steffen Dereich
Westfälische Wilhelms-Universität
Institut für Mathematische Statistik
Einsteinstr. 62
48149 Münster, Germany
E-mail: steffen.dereich@wwu.de

Mikhail Lifshits
St. Petersburg State University
Department of Mathematics and Mechanics
Bibliotechnaya pl., 2
198504 Stary Peterhof, Russia
and
MAI, Linköping University
58183 Linköping, Sweden
E-mail: mikhail@lifshits.org

Received on 14.1.2013;
revised version on 11.5.2013